The Degraded Poisson Wiretap Channel

Amine Laourine and Aaron B. Wagner

arXiv:1004.0269v1 [cs.IT] 2 Apr 2010

Abstract

Providing security guarantees for wireless communication is critically important for today's applications. While previous work in this area has concentrated on radio frequency (RF) channels, providing security guarantees for RF channels is inherently difficult because they are prone to rapid variations due small scale fading. Wireless optical communication, on the other hand, is inherently more secure than RF communication due to the intrinsic aspects of the signal propagation in the optical and near-optical frequency range. In this paper, secure communication over wireless optical links is examined by studying the secrecy capacity of a direct detection system. For the degraded Poisson wiretap channel, a closed-form expression of the secrecy capacity is given. A complete characterization of the general rate-equivocation region is also presented. For achievability, an optimal code is explicitly constructed by using the structured code designed by Wyner for the Poisson channel. The converse is proved in two different ways: the first method relies only on simple properties of the conditional expectation and basic information theoretical inequalities, whereas the second method hinges on the recent link established between minimum mean square estimation and mutual information in Poisson channels.

Index Terms

Information-theoretic security, wiretap channel, Poisson channel, direct detection optical communications.

Amine Laourine and Aaron B. Wagner are with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA. (Email: al496@cornell.edu, wagner@ece.cornell.edu.). Part of this paper will be presented in the 2010 IEEE International Symposium on Information Theory (ISIT 2010).

I. Introduction

PROTECTING information flow from unauthorized access is of vital importance for today's applications. The concept of information secrecy however is not new and dates back to the pioneering work of Shannon [1]. Shannon considered a strong notion of secrecy requiring essentially that the eavesdropper's received signal be independent of the communicated message. Under this constraint Shannon showed that the transmitter must share a private key with the legitimate receiver whose entropy is at least as large as the message's entropy. This negative result lead to the development of modern cryptographic systems whose objective is not to provide strong secrecy but rather to make security breaches computationally prohibitive. These encryption algorithms are employed today in most systems where secrecy is required. This wide deployment, however, should not obscure the fact that these methods do not offer strong security guarantees. Information-theoretical security can provide such guarantees and as such has attracted considerable attention lately.

The pessimistic result of Shannon is due to the implicit assumption that the eavesdropper observes the same signal as the legitimate receiver. In his celebrated paper [2], Wyner challenged this assumption by introducing the wiretap channel, a channel in which an eavesdropper observes a degraded version of the signal received by the legitimate receiver. Wyner also considered a different notion of secrecy which requires that asymptotically the equivocation of the eavesdropper about the transmitted message should converge to the message's entropy. With this new framework, and for discrete memoryless channels (DMCs), Wyner gave a complete description of the tradeoff between the information rate at the legitimate receiver and the information leaked to the eavesdropper. Wyner showed that for a DMC there exists an intrinsic quantity called the *secrecy capacity* which gives the maximum information rate that can be reliably transmitted to the legitimate receiver with zero leakage of information to the eavesdropper. Since then, the wiretap channel has become one of the core topics in information theoretic security and several results are available today. Csiszar and Korner [3] extended Wyner's study to nondegraded DMCs. The degraded Gaussian wiretap channel was studied in [4]. Multiple input multiple output (MIMO) Gaussian wiretap channels have been the subject of an extensive interest lately [5]-[7]. Wiretap channels in the presence of fading have been also investigated in [8]-[9] (and the references therein).

The results of [3] show that a non-zero secrecy capacity requires that the legitimate receiver's channel be less noisy

than the eavesdropper's. For RF channels, however, this is difficult to guarantee in practice due to the possibility of multipath fading. Indeed, even if the legitimate receiver is closer to the transmitter than the eavesdropper is, the legitimate receiver may still have a weaker channel due to fading. Moreover, since the fading state is a sensitive function of the position of the receivers, it is difficult to predict the degree of fading experienced by the legitimate receiver and the eavesdropper given imperfect information about their locations.

One possible solution to this problem is to use optical or near-optical frequencies instead of RF. For optical wireless systems, the detector is usually multiple orders of magnitude larger than the wavelength of the transmitted beam, which provides natural immunity against multipath fading via spatial diversity [10]. This immunity makes predictions about the quality of the legitimate receiver's and the eavesdropper's signal based on their position more accurate. In fact, with the multipath problem gone, the only major channel impairment remaining is the pathloss which can be safely assumed to be higher for the eavesdropper if the legitimate receiver can guarantee that he is closer to the transmitter.

Another advantage of optical communications over RF is that the transmitted signal is highly directional, making interception by a malevolent third party more difficult. This should be contrasted with the relatively-omnidirectional nature of RF transmissions, for which the signal is broadcasted over a wide angle. Yet another advantage of wireless optical communication is the spatial confinement of the transmitted optical signal. Indeed, at optical wavelengths, the transmitted signal is absorbed by the atmosphere and beyond a certain range it becomes undetectable. This is desirable from a security standpoint as an eavesdropper located beyond this range is literally kept in the dark.

All of these features give wireless optical communications a clear advantage for security. This technology is already being deployed in the form of infrared communications [11]-[12], and ultraviolet (UV) systems are currently under development [13]-[14]. However, despite the numerous benefits that this technology offers, coding is still needed for those scenarios in which the channel itself does not provide absolute secrecy, such as when the beam of light is reflected or scattered by solid objects, dust or water droplets [15]. Although in this case the signal has been degraded, an eavesdropper could still gain valuable information.

We examine the fundamental limits of coding for secure communication over optical channels by studying the secrecy capacity of the Poisson channel, a common model for direct detection optical communications systems. In

such systems the transmitter sends information by modulating the intensity of an optical signal while the receiver observes the arrival moments of individual photons. The capacity of this channel has been determined under peak power constraint on the transmitted optical power by Kabanov [16] and under both average and peak power constraint by Davis [17]. Wyner [18] derived the reliability function of this channel for all rates below capacity and constructed exponentially optimal codes. Multiple-access Poisson channels were studied in [19]-[20] whereas broadcast Poisson channels were considered in [21] and [22]. The capacity of the Poisson channel has been also investigated in the presence of fading [23].

We study in this paper the degraded Poisson wiretap channel. The legitimate receiver observes a doubly stochastic Poisson process with instantaneous rate $A_y X_t + \lambda_y$ where $\{X_t, 0 \leq t \leq T\}$ is the signal transmitted. The eavesdropper's observation is also a doubly stochastic Poisson process with instantaneous rate $A_z X_t + \lambda_z$. For degradedness we assume that[1] $A_y \geq A_z$ and $\lambda_y \leq \frac{A_y}{A_z}\lambda_z$. In Theorem 1 we provide a closed form expression of the secrecy capacity as a function of the parameters $(A_u, \lambda_u),\ u \in \{y, z\}$. This result is further extended by Theorem 5 which gives a full characterization of the rate equivocation region.

Our achievability proof uses stochastic encoding as well as the structured codes constructed by Wyner for the Poisson channel [18]. As for the converse, we will see that the infinite bandwidth nature of the Poisson channel makes it possible to prove the converse using only simple properties of the conditional expectation combined with basic information theoretical inequalities. This is to be contrasted with the converse of the (finite bandwidth) Gaussian channel which is proved using the entropy power inequality (EPI). As an illustration for the basic ideas that underpin the converse for the Poisson channel, we will start by considering here the more familiar infinite bandwidth Gaussian channel and we will see also that for this channel the proof of the converse simplifies considerably.

For this purpose, consider the continuous time Gaussian wiretap channel with bandwidth B (later we will let B tend to infinity) and with a power constraint P. This continuous time channel is equivalent to $2B$ uses per second of the discrete time Gaussian channel depicted in Fig. 1. The input signal is power constrained, i.e., $\mathbb{E}[X^2] \leq P$, the legitimate receiver observes $Y = X + W_1$ and the eavesdropper receives $Z = Y + W_2 = X + W_1 + W_2$, where $W_i \sim \mathcal{N}(0, N_i B)$ and $W_1 \perp\!\!\!\perp W_2$.

[1]These conditions were shown to be sufficient for degradedness in [21]. The argument is reproduced in Lemma 1 below.

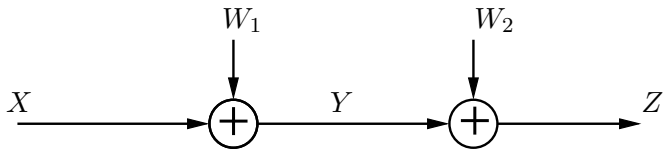

Fig. 1. The discrete time Gaussian wiretap channel

Define $\tilde{N}$ by $\frac{1}{\tilde{N}} = \frac{1}{N_1} - \frac{1}{N_1+N_2}$ and observe that

$$\frac{\tilde{N}}{N_1}Y = X + \frac{N_1}{N_2}Z + \tilde{W}, \tag{1}$$

where $\tilde{W} = W_1 - \frac{N_1}{N_2}W_2$. It is easy to see that $\tilde{W} \sim \mathcal{N}(0, \tilde{N}B)$ and $\mathbb{E}[\tilde{W}(W_1+W_2)] = 0$, it follows therefore that $\tilde{W} \perp\!\!\!\perp W_1 + W_2$ (since they are jointly Gaussian). For the discrete time Gaussian wiretap channel, it is known that the secrecy capacity is given by $\max_{p_X}(I(X;Y) - I(X;Z))$. For the continuous time channel counterpart with bandwidth B, the secrecy capacity becomes $C_s^B = 2B\max_{p_X}(I(X;Y) - I(X;Z))$.

In [4], using the celebrated EPI, a closed form expression for the secrecy capacity of the discrete time Gaussian wiretap channel was obtained. In just a few steps, we will see that the secrecy capacity of the infinite bandwidth ($B \to \infty$) Gaussian wiretap channel can be found much more simply. Starting with (1) we obtain the following sequence of inequalities

$$\begin{aligned}
I(X;Y) &= I(X; X + \frac{N_1}{N_2}Z + \tilde{W}) \\
&\overset{(a)}{\leq} I(X; X + \tilde{W}, Z) \\
&\overset{(b)}{=} h(X + \tilde{W}, Z) - h(X + \tilde{W}, Z|X) \\
&\overset{(c)}{\leq} h(X + \tilde{W}) + h(Z) - h(X + \tilde{W}, Z|X) \\
&\overset{(d)}{=} h(X + \tilde{W}) + h(Z) - h(X + \tilde{W}|X) - h(Z|X) \\
&\overset{(e)}{=} I(X; X + \tilde{W}) + I(X; Z),
\end{aligned}$$

Inequality (a) follows from the data processing inequality, equalities in (b) and (e) are standard information theory identities, (c) follows from the independence bound on entropy and finally (d) holds because $X + \tilde{W}$ and $X+W_1+W_2$ are conditionally independent given X (i.e., $(X+\tilde{W}) \perp\!\!\!\perp (X+W_1+W_2)|X$). Basically, the key identity needed to go from (a) to (e) is the following: if $Y_1 \perp\!\!\!\perp Y_2|X$, then we have $I(X;Y_1,Y_2) \leq I(X;Y_1) + I(X;Y_2)$. A

proof of this simple inequality in a more general setting will be given later and will be used in part of the converse for the Poisson channel. Going back to the Gaussian problem, we see that

$$C_s^B = 2B \max_{P_X}(I(X;Y) - I(X;Z)) \leq 2B \max_{P_X} I(X; X + \tilde{W}) = B \ln(1 + \frac{P}{B\tilde{N}}). \tag{2}$$

For a fixed bandwidth B, this last inequality is not tight. Now letting $B \to \infty$ we obtain

$$C_s^\infty \triangleq \lim_{B\to\infty} C_s^B \leq \frac{P}{\tilde{N}} = \frac{P}{N_1} - \frac{P}{N_1 + N_2}. \tag{3}$$

However, since $\max_{p_X}(I(X;Y) - I(X;Z)) \geq \max_{p_X} I(X;Y) - \max_{p_X} I(X;Z)$, we also have that

$$C_s^\infty \geq \lim_{B\to\infty} \left(B \ln(1 + \frac{P}{BN_1}) - B \ln(1 + \frac{P}{B(N_1 + N_2)}) \right) = \frac{P}{N_1} - \frac{P}{N_1 + N_2}. \tag{4}$$

It follows that $C_s^\infty = \frac{P}{N_1} - \frac{P}{N_1+N_2}$.

This remarkably simple approach will be particularly useful for the Poisson channel. More specifically, when $\frac{\lambda_y}{A_y} = \frac{\lambda_z}{A_z}$, the eavesdropper's signal Z is a thinned version of the legitimate receiver's signal Y, i.e., [2] $Y = Z + \tilde{Z}$ where $Z \perp\!\!\!\perp \tilde{Z}|X$, the approach above gives that $I(X;Y) - I(X;Z) \leq I(X;\tilde{Z})$. Since $\tilde{Z}$ is itself a doubly stochastic Poisson process, the mutual information $I(X;\tilde{Z})$ can be maximized using the martingale techniques of Kabanov [16] and an (achievable) upperbound can be obtained on $I(X;Y) - I(X;Z)$. When $\frac{\lambda_y}{A_y} < \frac{\lambda_z}{A_z}$, a different bounding technique using only simple properties of the conditional expectation will be devised.

Although no "sophisticated" tools are required to prove the converse, we show in the appendix that using some new results in information theory an alternative proof can be provided. This different proof hinges on the link that has been established between the mutual information (MI) and the minimum mean square estimation (MMSE) in Poisson channels [24]. It is worth noting at this point that the link between the MI and the MMSE in the Gaussian setting [25] has been also used recently for different Gaussian wiretap channels [26], [27].

One of the distinctive aspects in this paper is that we do not resort to the Δ-discretization method introduced by Wyner [18]. This method was used to approximate the Poisson channel by a binary DMC thereby allowing the transposition of the widely known results for DMCs to the Poisson channel. This technique leads to extensive computations, especially when we are interested in the secrecy capacity as there are now two conflicting objectives

[2]The time dependence has been dropped to ease the notations. Refer to the converse part of the paper for a mathematically precise statement.

involved, the maximization of the information rate at the legitimate receiver and the minimization of the information leakage at the eavesdropper. We circumvent the use of this method by using the techniques described above.

The rest of this paper is organized as follows. The next section describes the setup of the problem and presents the main result of this paper as well as some interpretations of the obtained result. The proof of the achievability of the secrecy capacity is given in Section III and the proof of the converse is presented in Section IV. In Section V we extend the main result of the paper by giving a complete characterization of the rate-equivocation region. Finally, in section VI, some possible future directions are discussed.

II. Problem and Result Statement

The input process to the Poisson channel is a waveform denoted by $X_0^T \triangleq \{X_t, 0 \leq t \leq T\}$ satisfying $X_t \geq 0$ for all t. We further assume that the input process is peak power limited, i.e., $X_t \leq 1$ for all t. The received signal at the legitimate receiver Y_0^T is a doubly stochastic Poisson process with instantaneous rate $A_y X_t + \lambda_y$, i.e., given X_0^T the stochastic process Y_0^T has independent increments with $Y_0 = 0$ and for $0 \leq s \leq t \leq T$ we have

$$\Pr(Y_t - Y_s = k | X_0^T) = \frac{1}{k!}\Upsilon^k(s,t)e^{-\Upsilon(s,t)}, \ k \in \mathbb{N},$$

where

$$\Upsilon(s,t) = \int_s^t (A_y X_\tau + \lambda_y) d\tau.$$

The parameter $A_y > 0$ accounts for possible signal attenuation at the receiver. The parameter $\lambda_y \geq 0$ is the dark current intensity which results from background noise and bears no information on the input process X_0^T. Similarly the output process of the eavesdropper Z_0^T is a doubly stochastic Poisson process with instantaneous rate $A_z X_t + \lambda_z$.

In this paper, the space of doubly stochastic Poisson processes on the interval $[0, T]$ will be denoted by $\mathcal{P}(T)$. Following the notation used in [24] the output process of the Poisson channel in the interval $[0, T]$ with instantaneous rate $\alpha X_t + \lambda$ will be denoted by $\mathcal{P}_0^T(\alpha X_0^T + \lambda)$. We use $\langle X_t \rangle_s$ to designate $E[X_t | \mathcal{P}_0^s(X_0^s)]$, as such $\langle X_t \rangle_t$ refers to the causal conditional mean estimate and $\langle X_t \rangle_T$ to the noncausal one.

All stochastic processes considered in this paper are defined on a common measurable space $(\Omega, \mathcal{F})$. We use $\mathcal{F}_\xi^s$ to denote the internal history generated by the process ξ_0^s.

In this paper we are interested in the degraded Poisson wiretap channel. Lapidoth et al. [21] gave conditions on the parameters $(A_u, \lambda_u),\ u \in \{y, z\}$ for stochastic degradedness. These conditions are presented in the following lemma. In order to prepare for the results to come we will also briefly go over the proof of this lemma.

Lemma 1 (Lapidoth, Telatar and Urbanke [21]). *The eavesdropper's channel is stochastically degraded with respect to the legitimate receiver's channel, if*

$$A_y \geq A_z, \tag{5}$$

and

$$\lambda_y \leq \frac{A_y}{A_z}\lambda_z. \tag{6}$$

Proof: Let $\tilde{Y}_0^T$ (cf. Fig. 2) be the process defined as follows

$$\tilde{Y}_t = Y_t + H_t,\ t \in [0, T], \tag{7}$$

where H_t is a homogeneous Poisson process with rate $\tilde{\lambda} = \frac{A_y}{A_z}\lambda_z - \lambda_y$ (note that $\tilde{\lambda} \geq 0$ by (6)) independent of (X_0^T, Y_0^T). It follows that $\tilde{Y}_0^T$ is a doubly stochastic Poisson process with instantaneous rate $A_y X_t + \lambda_y + \tilde{\lambda} = A_y X_t + \frac{A_y}{A_z}\lambda_z$. The process Z_0^T is then obtained from $\tilde{Y}_0^T$ by thinning with erasure probability $1 - \frac{A_z}{A_y}$ (note that because of (5) this quantity is ≥ 0). ■

In the rest of this paper we will assume that at least one of the inequalities (5) or (6) is strict. Note that this assumption can be made without losing generality for if there was an equality in (5) and (6) then the legitimate receiver's channel and the eavesdropper's channel will be identical and the secrecy capacity will be zero.

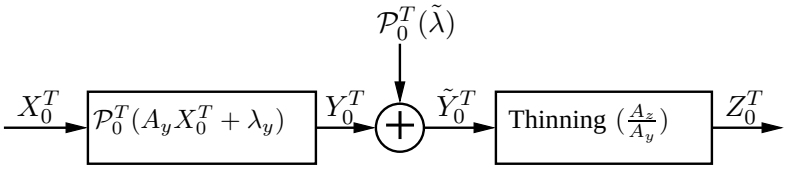

Fig. 2. The degraded Poisson wiretap channel

We move now to the description of the information transmission aspect of the problem. The transmitter wishes to communicate a message U uniformly distributed on $\mathcal{U} = \{1, ..., M\}$. An (M, T) code (E_T, D_T) for the Poisson wiretap channel is a stochastic encoder E_T that maps a message U to a waveform X_0^T which satisfies the peak

power constraint and a decoder $D_T : \mathcal{P}(T) \to \mathcal{U}$. The transmission rate of this code is

$$R = \frac{H(U)}{T} = \frac{1}{T} \ln M.$$

The average probability of error at the legitimate receiver is

$$P_e = \frac{1}{M} \sum_{m=1}^{M} \Pr\left(D_T(Y_0^T) \neq m | U = m\right). \tag{8}$$

The level of secrecy in this paper is measured by $\frac{1}{T} I(U; Z_0^T)$. This normalized mutual information quantifies the amount of information about the message U leaked to the eavesdropper. As such our goal is to make this quantity as small as possible.

Definition A secrecy rate R_s is said to be *achievable* for the Poisson wiretap channel if for all $\epsilon > 0$ and all sufficiently large T, there exists an (M, T) code such that

$$\frac{\ln M}{T} \geq R_s - \epsilon$$

$$P_e \leq \epsilon$$

$$\frac{1}{T} I(U; Z_0^T) \leq \epsilon \tag{9}$$

The supremum of achievable secrecy rates will be called the *secrecy capacity*. The main result of this paper is the following.

Theorem 1. *The secrecy capacity of the degraded Poisson wiretap channel is given by*[3]

$$C_s = \alpha^*(A_y - A_z) + \ln\left(\frac{\lambda_y^{\lambda_y}}{\lambda_z^{\lambda_z}}\right) + \ln\left(\frac{(A_z\alpha^* + \lambda_z)^{\lambda_z}}{(A_y\alpha^* + \lambda_y)^{\lambda_y}}\right), \tag{10}$$

where α^* is the unique solution in $[0, 1]$ to the following equation

$$\frac{(A_y\alpha^* + \lambda_y)^{A_y}}{(A_z\alpha^* + \lambda_z)^{A_z}} = e^{A_z - A_y} \frac{(A_y + \lambda_y)^{A_y + \lambda_y}}{(A_z + \lambda_z)^{A_z + \lambda_z}} \frac{\lambda_z^{\lambda_z}}{\lambda_y^{\lambda_y}}. \tag{11}$$

This result assumes that (A_z, λ_z) is known to the transmitter. Yet it follows that C_s is an achievable rate with perfect secrecy even if the eavesdropper observes $\mathcal{P}_0^T(A_z' X_0^T + \lambda_z')$, where A_z' and λ_z' are unknown but satisfy $A_z' \leq A_z$ and $\lambda_z' \geq \frac{A_z'}{A_z}\lambda_z$. Thus, only one-sided estimates of A_z and λ_z are needed. In practice, an upper bound on A_z could be provided by guaranteeing that any potential eavesdropper is more than a certain distance away from

[3]If $\lambda = 0$, the convention is that $0^0 = 1$.

the transmitter. A lower bound on the dark current λ_z could be provided using ambient noise measurements and the known physical limitations of existing receivers.

Worst case scenario: A particularly insightful case is when $\frac{\lambda_y}{A_y} = \frac{\lambda_z}{A_z} = \sigma$. This situation happens when the eavesdropper observes a thinned version of the signal of the legitimate receiver, i.e., $H_t \equiv 0$ in (7). In this case, after some algebraic manipulations, we obtain that

$$\alpha^* = \frac{(1+\sigma)^{1+\sigma}}{e\sigma^{\sigma}} - \sigma, \tag{12}$$

and the secrecy capacity reduces to

$$C_s = (\lambda_y - \lambda_z)\left(\frac{1}{e}\left(1+\frac{1}{\sigma}\right)^{1+\sigma} - (1+\sigma)\ln\left(1+\frac{1}{\sigma}\right)\right). \tag{13}$$

This is saying that the secrecy capacity is the difference between the capacity of the main channel (the channel between the transmitter and the legitimate receiver) and the capacity of the eavesdropper's channel. For instance, in the special case when there is no dark current $\lambda_y = \lambda_z = 0$, we find that $\alpha^* = \frac{1}{e}$ and the secrecy capacity reduces to

$$C_s = \frac{A_y - A_z}{e}. \tag{14}$$

For a degraded DMC, Wyner [2] showed that the secrecy capacity is equal to $\max_{p_X}(I(X;Y) - I(X;Z))$. Hence the following inequality is always satisfied

$$\text{Secrecy Capacity} \geq C_M - C_W,$$

where C_M is the capacity of the main channel and C_W is the capacity of the eavesdropper's channel. As shown in [28], there is equality in the inequality above if there is an input probability distribution p_X that maximizes simultaneously $I(X;Y)$ and $I(X;Z)$. This is exactly what is happening here, when $\frac{\lambda_y}{A_y} = \frac{\lambda_z}{A_z}$ the mutual information $I(X_0^T;Y_0^T)$ and $I(X_0^T;Z_0^T)$ are both maximized by letting the input X_0^T cycle infinitely fast between its extreme values, i.e., the peak power 1 and 0 with $\Pr(X_t = 1) = 1 - \Pr(X_t = 0) = \alpha^* = \frac{(1+\sigma)^{1+\sigma}}{e\sigma^{\sigma}} - \sigma$.

Before we proceed further with the presentation of the problem considered in this paper, we give a lemma that will prove to be useful in the proofs of the achievability and the converse, a proof of this result can be found for instance in [16].

Lemma 2. *The mutual information between the input X_0^T and the output $\mathcal{P}_0^T(\alpha X_0^T + \lambda)$ can be upper bounded as follows*[4]

$$I(X_0^T; \mathcal{P}_0^T(\alpha X_0^T + \lambda)) \leq \int_0^T (\mathbb{E}[\vartheta(X_t)] - \vartheta(\mathbb{E}[X_t]))dt, \tag{15}$$

where $\vartheta(x) = (\alpha x + \lambda)\ln(\alpha x + \lambda)$.

III. Achievability of C_s

Our achievability proof relies on the structured codes that were designed for the Poisson channel by Wyner [18]. Before delving into the details of the proof, we will briefly describe the code construction and the properties inherited by this code.

Wyner codes $\mathcal{W}(T, M, k)$: Let T, M and k be given, and construct an $M \times \binom{M}{k}$ binary matrix $\mathcal{X}$ as follows. The columns of $\mathcal{X}$ are the $\binom{M}{k}$ binary M-vectors with exactly k ones and $M-k$ zeros. Now partition the interval $[0, T]$ into $\binom{M}{k}$ subintervals of equal length $\varpi_T \triangleq \frac{T}{\binom{M}{k}}$ and construct M waveforms $\{X_0^T(m)\}_{m=1}^M$ as follows

$$X_t(m) = \mathcal{X}(m, n), t \in ((n-1)\varpi_T, n\varpi_T],\ n = 1, ..., \binom{M}{k}. \tag{16}$$

For $\alpha = \frac{k}{M}$ fixed, these codes satisfy

$$\frac{1}{T}\mu\{t : X_t(m) = 1\} = \alpha, \quad \text{for all } m, \tag{17}$$

with μ being the Lebesgue measure. If moreover $M = e^{RT}$, for $T >> 1$, Wyner showed that for $m \neq m'$

$$\frac{1}{T}\mu\{t : X_t(m) = 1, X_t(m') = 0\} \approx \alpha(1-\alpha). \tag{18}$$

As such for T large enough the codewords $\{X_0^T(m)\}_{m=1}^M$ will behave as if they were chosen independently.

After this brief overview of Wyner codes we are in a position to state the achievability theorem and prove it.

Theorem 2. *Any secrecy rate $R_s < C_s$ is achievable.*

Proof: Let $\epsilon > 0$ be arbitrary and let $R_s = C_s - \epsilon$. Define

$$R_u = \alpha^*(A_u + \lambda_u)\ln(A_u + \lambda_u) + (1-\alpha^*)\lambda_u \ln \lambda_u$$
$$- (A_u\alpha^* + \lambda_u)\ln(A_u\alpha^* + \lambda_u),\ u \in \{y, z\}. \tag{19}$$

[4]Note that some authors use the function $\vartheta(x) = (\alpha x + \lambda)\ln(\alpha x + \lambda) - \lambda \ln \lambda$ instead but the constant term $\lambda \ln \lambda$ cancels out here.

After few algebraic manipulations, we can show that

$$C_s = R_y - R_z.$$

Given these parameters, the encoder-decoder pair considered here works as follows.

Encoding: Let $M = e^{R_s T}$ and let U be uniformly distributed on $\mathcal{U} = \{1, ..., M\}$. Define $M_y = e^{(R_y - \frac{3}{2}\epsilon)T}$ and following the steps described above construct a code[5] $\mathcal{C} = \mathcal{W}(T, M_y, \alpha^* M_y)$. Partition this code arbitrarily into M smaller subcodes, i.e., $\mathcal{C} = \cup_{i=1}^{M} \mathcal{C}_i$. The cardinality of each each subcode $\mathcal{C}_i$ will be equal to $M_z = \frac{M_y}{M} = e^{(R_z - \frac{\epsilon}{2})T}$.

The encoder works as follows, when the message $U = m$ is chosen, the codeword X_0^T is selected uniformly randomly from $\mathcal{C}_m$.

Decoding: The decoder considered here is the maximum likelihood decoder constructed by Wyner [18]. After observing Y_0^T, the decoder at the legitimate receiver computes the following metric

$$\Psi_m = \int_{S_m} dY_t, \tag{20}$$

where $S_m = \{t \in [0, T] : \ X_t(m) = 1\}$. Then $D_T(Y_0^T) = m$ if m maximizes Ψ_m, with ties resolved in favor of the smallest m.

Analysis of P_e: The fact that $P_e \to 0$, follows simply from the fact that Wyner codes with the peak power 1 and average power α^* are capacity achieving.

Analysis of $\frac{1}{T} I(U; Z_0^T)$:

Notice first that for each m, the waveform $X_0^T(m)$ is piecewise constant. It follows that a sufficient statistic for making a decision is the number of arrivals during each subinterval $((n-1)\varpi_T, n\varpi_T]$, i.e., $Z_n = Z_{n\varpi_T} - Z_{(n-1)\varpi_T}$, $n = 1, ..., N_y$ with $N_y = \binom{M_y}{\alpha^* M_y}$. Consequently,

$$I(X_0^T; Z_0^T) = I(\mathbf{X}; \mathbf{Z})$$

$$I(U; Z_0^T) = I(U; \mathbf{Z}) \tag{21}$$

where $\mathbf{X} = [X_1, ..., X_{N_y}]$, $X_i = 0$ or 1 depending on the choice of the codeword and $\mathbf{Z} = [Z_1, ..., Z_{N_y}]$. The equalities above follows from the fact that $\mathbf{Z}$ is a sufficient statistic.

[5]Note that even if α^* is not a rational number, it can be approximated arbitrary close by rationals.

As a result of Lemma 2, we have

$$\frac{1}{T}I(X_0^T;Z_0^T) \leq \frac{1}{T}\int_0^T (\mathbb{E}[\phi_z(X_t)] - \phi_z(\mathbb{E}[X_t]))dt, \tag{22}$$

where $\phi_z(x) = (A_z x + \lambda_z)\ln(A_z x + \lambda_z)$. Because of the uniform choice in the encoding scheme and in view of (17) we must have that $\Pr[X_t = 1] = 1 - \Pr[X_t = 0] = \alpha^*$, hence we have

$$\begin{aligned}\mathbb{E}[\phi_z(X_t)] &= \alpha^*\phi_z(1) + (1-\alpha^*)\phi_z(0) \\ &= \alpha^*(A_z + \lambda_z)\ln(A_z + \lambda_z) + (1-\alpha^*)\lambda_z \ln \lambda_z. \end{aligned}\tag{23}$$

and

$$\phi_z(\mathbb{E}[X_t]) = \phi_z(\alpha^*) = (A_z\alpha^* + \lambda_z)\ln(A_z\alpha^* + \lambda_z). \tag{24}$$

Consequently, we deduce that

$$\frac{1}{T}I(\mathbf{X};\mathbf{Z}) = \frac{1}{T}I(X_0^T;Z_0^T) \leq R_z. \tag{25}$$

Notice that every subcode $\mathcal{C}_m$ can be viewed as a code for the eavesdropper's channel with M_z codewords and uniform prior distribution. Define δ_m to be the probability of error for code $\mathcal{C}_m$ ($1 \leq m \leq M$) with the (optimal) decoder described above and let $\delta = \frac{1}{M}\sum_{m=1}^{M}\delta_m$. From Fano's inequality we have

$$H(\mathbf{X}|\mathbf{Z}, U = m) \leq H(\delta_m) + \delta_m \ln M_z, \tag{26}$$

where $H(p) = -p\ln(p) - (1-p)\ln(1-p)$ is the binary entropy.

Since the codewords are uniformly distributed in each subcode, we deduce that $H(\mathbf{X}|U = m) = \ln M_z$. We conclude therefore that

$$\begin{aligned} I(\mathbf{X};\mathbf{Z}|U=m) &= H(\mathbf{X}|U=m) - H(\mathbf{X}|\mathbf{Z},U=m) \\ &\geq \ln M_z - (H(\delta_m) + \delta_m \ln M_z). \end{aligned}\tag{27}$$

Averaging over U and by using the concavity of $H(\cdot)$ we find that

$$I(\mathbf{X};\mathbf{Z}|U) \geq \ln M_z - (H(\delta) + \delta \ln M_z). \tag{28}$$

Notice also that $U \to \mathbf{X} \to \mathbf{Z}$ form a Markov chain, i.e.,

$$\begin{aligned}\frac{1}{T}I(U;\mathbf{Z}) &= \frac{1}{T}I(U,\mathbf{X};\mathbf{Z}) - \frac{1}{T}I(\mathbf{X};\mathbf{Z}|U) \\ &= \frac{1}{T}I(\mathbf{X};\mathbf{Z}) - \frac{1}{T}I(\mathbf{X};\mathbf{Z}|U) \end{aligned} \tag{29}$$

Combined with the last inequality this implies that

$$\frac{1}{T}I(U;\mathbf{Z}) \leq \frac{1}{T}I(\mathbf{X};\mathbf{Z}) - \frac{1}{T}\ln M_z + \frac{1}{T}(H(\delta) + \delta \ln M_z). \tag{30}$$

Inequalities (25) and (30) result in the following

$$\frac{1}{T}I(U;\mathbf{Z}) \leq R_z - \frac{1}{T}[\ln M_z - (H(\delta) + \delta \ln M_z)]. \tag{31}$$

As $M_z = e^{(R_z - \frac{\epsilon}{2})T}$, this gives

$$\frac{1}{T}I(U;Z_0^T) = \frac{1}{T}I(U;\mathbf{Z}) \leq \frac{\epsilon}{2} + \frac{1}{T}H(\delta) + \delta(R_z - \frac{\epsilon}{2}). \tag{32}$$

By the code construction described above, the codewords of every subcode $\mathcal{C}_m$ ($1 \leq m \leq M$) satisfy (17) and (18) (with α replaced by α^*). These two conditions dictate the pairwise error probability of the codewords in $\mathcal{C}_m$ [18]. Since the overall error probability of the code $\mathcal{C}_m$ is governed by the pairwise error probability [18], it follows that every subcode $\mathcal{C}_m$ is capacity achieving for the eavesdropper's channel and as such δ_m for $m = 1, ..., M$ can be made arbitrarily small. Hence, by choosing T large enough, we can enforce that $\frac{1}{T}H(\delta) + \delta(R_z - \frac{\epsilon}{2}) \leq \frac{\epsilon}{2}$. The previous inequality shows therefore that $\frac{1}{T}I(U;Z_0^T) \leq \epsilon$ and the desired secrecy condition is satisfied.

This shows that any secrecy rate $R_s < C_s$ can be achieved and completes the achievability proof. ■

IV. The Converse for the secrecy capacity

Before delving into the details of the converse we need the following technical lemma due to Wyner [29].

Lemma 3 (Wyner [29]). *If $\Gamma : \Omega \to F$ is a random variable such that F is a finite set and $\Lambda_0^T = \{\Lambda_t, 0 \leq t \leq T\}$ is a given stochastic process, then we have*

$$I(\Gamma;\Lambda_0^T) = H(\Gamma) - H(\Gamma|\Lambda_0^T), \tag{33}$$

where $H(\Gamma)$ is the usual entropy for discrete random variables and

$$H(\Gamma|\Lambda_0^T) = -\mathbb{E}\left[\sum_{\gamma \in F} \Pr[\Gamma = \gamma|\mathcal{F}_\Lambda^T] \ln \Pr[\Gamma = \gamma|\mathcal{F}_\Lambda^T]\right]. \tag{34}$$

This lemma is standard when all the random variables have discrete alphabets however this extension is needed in this paper since we are dealing with continuous time stochastic processes.

The converse theorem will be proved through a sequence of Lemmas. The first one gives an inequality that must satisfied by every encoder-decoder pair (E_T, D_T).

Lemma 4. *For every (M,T) code with rate $R = \frac{\ln M}{T}$ we have*

$$R \leq \frac{1}{T(1-P_e)} \left(I(X_0^T; Y_0^T | Z_0^T) + I(U; Z_0^T) + H(P_e) \right). \tag{35}$$

Proof: Let $\hat{U} = D_T(Y_0^T)$ denote the output of the decoder at the legitimate receiver, so that $P_e = \Pr(U \neq \hat{U})$. We then have the following sequence of identities

$$\begin{aligned} RT = \ln M = H(U) &\overset{(a)}{=} H(U|Y_0^T) + I(U; Y_0^T) \\ &\overset{(b)}{\leq} H(U|\hat{U}) + I(U; Y_0^T) \\ &\overset{(c)}{\leq} H(P_e) + P_e \ln M + I(U; Y_0^T), \end{aligned} \tag{36}$$

the equality (a) follows from Wyner's lemma and the inequality (c) is an application of Fano's inequality. For the inequality (b), since $U \to Y_0^T \to \hat{U}$ is a Markov chain we deduce that[6] $I(U, Y_0^T) \geq I(U, \hat{U})$. Now by invoking Wyner's lemma again it follows that $H(U|Y_0^T) \leq H(U|\hat{U})$.

From Kolmogorov's formula (see Lemma 3.2 in [29]) we have[7]

$$I(U; Y_0^T, Z_0^T) = I(U; Y_0^T) + I(U; Z_0^T | Y_0^T), \tag{37}$$

since $U \to Y_0^T \to Z_0^T$ is a Markov chain we deduce that[8] $I(U; Z_0^T | Y_0^T) = 0$. By applying Kolmogorov's formula again we obtain

$$\begin{aligned} I(U; Y_0^T) = I(U; Y_0^T, Z_0^T) &= I(U; Z_0^T) + I(U; Y_0^T | Z_0^T) \\ &\leq I(U; Z_0^T) + I(X_0^T; Y_0^T | Z_0^T), \end{aligned} \tag{38}$$

where the last inequality follows from the fact that $U \to X_0^T \to Y_0^T \to Z_0^T$ form a Markov chain. Combining this last inequality with (c) and rearranging the terms yields the desired inequality. ■

[6]The data processing inequality extends to arbitrary random variables, see for instance Theorem 3.4 in [29].

[7]The definition of the conditional mutual information for arbitrary random variables can be found in [29].

[8]Refer to Lemma 3.1. in [29].

Lemma 5. *If* $\mathbb{E}\int_0^T |X_t \ln X_t| < \infty$*, then*

$$I(X_0^T; Y_0^T | Z_0^T) = I(X_0^T; Y_0^T) - I(X_0^T; Z_0^T) \tag{39}$$

Proof: Applying Kolmogorov's formula twice gives

$$\begin{aligned} I(X_0^T; Y_0^T, Z_0^T) &= I(X_0^T; Z_0^T) + I(X_0^T; Y_0^T | Z_0^T) \\ &= I(X_0^T; Y_0^T) + I(X_0^T; Z_0^T | Y_0^T). \end{aligned} \tag{40}$$

Since $X_0^T \to Y_0^T \to Z_0^T$ form a Markov chain, we have $I(X_0^T; Z_0^T | Y_0^T) = 0$. Consequently we deduce that

$$I(X_0^T; Y_0^T) = I(X_0^T; Z_0^T) + I(X_0^T; Y_0^T | Z_0^T). \tag{41}$$

The condition $\mathbb{E}\int_0^T |X_t \ln X_t| < \infty$ implies that $I(X_0^T; Z_0^T) < \infty$ and it follows that $I(X_0^T; Y_0^T) - I(X_0^T; Z_0^T) = I(X_0^T; Y_0^T | Z_0^T)$. ■

The goal of the upcoming lemmas is to prove that $I(X_0^T; Y_0^T | Z_0^T) \leq TC_s$, where C_s is given by (10). We first decompose $I(X_0^T; Y_0^T | Z_0^T)$ as follows

$$\begin{aligned} I(X_0^T; Y_0^T | Z_0^T) &= I(X_0^T; Y_0^T) - I(X_0^T; Z_0^T) \\ &= I(X_0^T; Y_0^T) - I(X_0^T; \tilde{Y}_0^T) \\ &\quad + I(X_0^T; \tilde{Y}_0^T) - I(X_0^T; Z_0^T), \end{aligned} \tag{42}$$

where $\tilde{Y}_0^T$ has been defined in (7). The next two lemmas will provide upper bounds on $I(X_0^T; Y_0^T) - I(X_0^T; \tilde{Y}_0^T)$ and $I(X_0^T; \tilde{Y}_0^T) - I(X_0^T; Z_0^T)$.

Lemma 6. *If* $\mathbb{E}\int_0^T |X_t \ln X_t| < \infty$*, then*

$$\begin{aligned} &I(X_0^T; Y_0^T) - I(X_0^T; \tilde{Y}_0^T) \leq \\ &\int_0^T \left(\frac{A_y}{A_z} (\phi_z(\mathbb{E}[X_t]) - \mathbb{E}[\phi_z(X_t)]) - (\phi_y(\mathbb{E}[X_t]) - \mathbb{E}[\phi_y(X_t)]) \right) dt, \end{aligned} \tag{43}$$

where $\phi_y(x) = (A_y x + \lambda_y)\ln(A_y x + \lambda_y)$ *and* $\phi_z(x)$ *has been defined above analogously.*

Proof: Note first that [16], [30]

$$I(X_0^T; Y_0^T) = \int_0^T \left(\mathbb{E}[\phi_y(X_t)] - \mathbb{E}[\phi_y(\mathbb{E}[X_t | \mathcal{F}_Y^t])] \right) dt, \tag{44}$$

and

$$I(X_0^T;\tilde{Y}_0^T) = \int_0^T \Big(\mathbb{E}[\chi(X_t)] - \mathbb{E}[\chi(\mathbb{E}[X_t|\mathcal{F}_{\tilde{Y}}^t])]\Big)\, dt, \tag{45}$$

where $\chi(x) = (A_y x + \frac{A_y}{A_z}\lambda_z)\ln(A_y x + \frac{A_y}{A_z}\lambda_z)$. Consequently, using the fact that $\chi(x) = \frac{A_y}{A_z}\phi_z(x) + \ln(\frac{A_y}{A_z})(A_y x + \frac{A_y}{A_z}\lambda_z)$ and after simplifications, we deduce the following

$$\begin{aligned} I(X_0^T;Y_0^T) - I(X_0^T;\tilde{Y}_0^T) = &\int_0^T \big(\mathbb{E}[\phi_y(X_t)] - \mathbb{E}[\phi_y(\mathbb{E}[X_t|\mathcal{F}_Y^t])]\big)\, dt \\ &- \frac{A_y}{A_z}\int_0^T \Big(\mathbb{E}[\phi_z(X_t)] - \mathbb{E}[\phi_z(\mathbb{E}[X_t|\mathcal{F}_{\tilde{Y}}^t])]\Big)\, dt. \end{aligned} \tag{46}$$

Recall that $\tilde{Y}_0^T = Y_0^T + H_0^T$, where H_0^T is a homogeneous Poisson process independent of (X_0^T, Y_0^T). Clearly, $\mathcal{F}_{\tilde{Y}}^t \subset \mathcal{F}_Y^t \vee \mathcal{F}_H^t$, with $\mathcal{F}_Y^t \vee \mathcal{F}_H^t = \sigma(\mathcal{F}_Y^t \cup \mathcal{F}_H^t)$ being the smallest sigma-field containing $\mathcal{F}_Y^t \cup \mathcal{F}_H^t$. From the independence of (X_0^T, Y_0^T) from H_0^T, using the law of redundant conditioning (see, e.g. [30, pp. 281-282]), we deduce that

$$\mathbb{E}[X_t|\mathcal{F}_Y^t \vee \mathcal{F}_H^t] = \mathbb{E}[X_t|\mathcal{F}_Y^t] \quad \text{a.s.} \tag{47}$$

We can now establish the following sequence of identities

$$\mathbb{E}[\phi_z(\mathbb{E}[X_t|\mathcal{F}_Y^t])] \overset{(a)}{=} \mathbb{E}[\phi_z(\mathbb{E}[X_t|\mathcal{F}_Y^t \vee \mathcal{F}_H^t])] \tag{48}$$

$$\overset{(b)}{=} \mathbb{E}[\mathbb{E}[\phi_z(\mathbb{E}[X_t|\mathcal{F}_Y^t \vee \mathcal{F}_H^t])|\mathcal{F}_{\tilde{Y}}^t]] \tag{49}$$

$$\overset{(c)}{\geq} \mathbb{E}[\phi_z(\mathbb{E}[\mathbb{E}[X_t|\mathcal{F}_Y^t \vee \mathcal{F}_H^t]|\mathcal{F}_{\tilde{Y}}^t])] \tag{50}$$

$$\overset{(d)}{=} \mathbb{E}[\phi_z(\mathbb{E}[X_t|\mathcal{F}_{\tilde{Y}}^t])], \tag{51}$$

where (a) follows from (47), (b) follows from the smoothing property of the conditional expectation, (c) from Jensen's inequality applied to the convex function $\phi_z(\cdot)$ and (d) from the fact that $\mathcal{F}_{\tilde{Y}}^t \subset \mathcal{F}_Y^t \vee \mathcal{F}_H^t$ and the smoothing property.

We deduce therefore that

$$\begin{aligned} I(X_0^T;Y_0^T) - I(X_0^T;\tilde{Y}_0^T) \leq &\int_0^T \big(\mathbb{E}[\phi_y(X_t)] - \mathbb{E}[\phi_y(\mathbb{E}[X_t|\mathcal{F}_Y^t])]\big)\, dt \\ &- \frac{A_y}{A_z}\int_0^T \big(\mathbb{E}[\phi_z(X_t)] - \mathbb{E}[\phi_z(\mathbb{E}[X_t|\mathcal{F}_Y^t])]\big)\, dt. \end{aligned} \tag{52}$$

A simple derivation shows that the function $\pi(x) = \phi_y(x) - \frac{A_y}{A_z}\phi_z(x)$ is convex as

$$\pi''(x) = \frac{A_y(\lambda_z A_y - \lambda_y A_z)}{(A_y x + \lambda_y)(A_z x + \lambda_z)} \geq 0. \tag{53}$$

Now invoking again Jensen's inequality we obtain that

$$\mathbb{E}[\pi(\mathbb{E}[X_t|\mathcal{F}_Y^t])] \geq \pi(\mathbb{E}[\mathbb{E}[X_t|\mathcal{F}_Y^t]]) = \pi(\mathbb{E}[X_t]). \tag{54}$$

Using this last inequality and after rearranging the terms we obtain the desired result, i.e.,

$$\begin{aligned} I(X_0^T; Y_0^T) - I(X_0^T; \tilde{Y}_0^T) \leq & \int_0^T (\mathbb{E}[\phi_y(X_t)] - \phi_y(\mathbb{E}[X_t]))\, dt \\ & - \frac{A_y}{A_z} \int_0^T (\mathbb{E}[\phi_z(X_t)] - \phi_z(\mathbb{E}[X_t]))\, dt. \end{aligned} \tag{55}$$

■

An alternative proof of this lemma using the link provided in [24] between the MMSE and the mutual information in Poisson channels is given in Appendix A.

Lemma 7. *If* $\mathbb{E}\int_0^T |X_t \ln X_t| < \infty$*, then*

$$\begin{aligned} & I(X_0^T; \tilde{Y}_0^T) - I(X_0^T; Z_0^T) \\ & \leq (\frac{A_y}{A_z} - 1) \int_0^T (\mathbb{E}[\phi_z(X_t)] - \phi_z(\mathbb{E}[X_t]))dt \end{aligned} \tag{56}$$

Proof: Recall that Z_0^T was obtained from $\tilde{Y}_0^T$ by thinning with erasure probability $1 - \frac{A_z}{A_y}$. Let the process $\tilde{Z}_0^T$ denote those points that were erased, hence we have that $\tilde{Z}_0^T$ is a doubly stochastic Poisson process with instantaneous rate $(A_y - A_z)X_t + (\frac{A_y}{A_z} - 1)\lambda_z$. Moreover Z_0^T and $\tilde{Z}_0^T$ are independent given X_0^T. We proceed with the proof of the lemma by showing that the following inequality holds

$$I(X_0^T; \tilde{Y}_0^T) - I(X_0^T; Z_0^T) \leq I(X_0^T; \tilde{Z}_0^T). \tag{57}$$

Indeed, notice first that $X_0^T \to (Z_0^T, \tilde{Z}_0^T) \to \tilde{Y}_0^T$ is a Markov chain, hence from the data processing inequality we deduce that

$$I(X_0^T; \tilde{Y}_0^T) = I(X_0^T; Z_0^T + \tilde{Z}_0^T) \leq I(X_0^T; Z_0^T, \tilde{Z}_0^T). \tag{58}$$

Consider now two partitions of Ω, $\mathcal{Q}_Z = \{A_i\}_{i=1}^{N_1} \subseteq \mathcal{F}_Z^T$ and $\mathcal{Q}_{\tilde{Z}} = \{B_j\}_{j=1}^{N_2} \subseteq \mathcal{F}_{\tilde{Z}}^T$. Define two discrete random variables D and $\tilde{D}$ on Ω as follows $D(\omega) = i$ if $\omega \in A_i$ and $\tilde{D}(\omega) = j$ if $\omega \in B_j$. The mutual information

$I(X_0^T; Z_0^T, \tilde{Z}_0^T)$ can be computed as [29]

$$I(X_0^T; Z_0^T, \tilde{Z}_0^T) = \sup_{\mathcal{Q}_z, \mathcal{Q}_{\tilde{z}}} I(X_0^T; D, \tilde{D}), \tag{59}$$

where the supremum is taken over all such partitions of Ω. We proceed to prove (57) as follows

$$\begin{aligned} I(X_0^T; D, \tilde{D}) &\overset{(a)}{=} H(D, \tilde{D}) - H(D, \tilde{D}|X_0^T) \\ &\overset{(b)}{\leq} H(D) + H(\tilde{D}) - H(D, \tilde{D}|X_0^T) \\ &\overset{(c)}{=} H(D) + H(\tilde{D}) - H(D|X_0^T) - H(\tilde{D}|X_0^T) \\ &\overset{(d)}{=} I(D; X_0^T) + I(\tilde{D}; X_0^T), \end{aligned} \tag{60}$$

where (a) follows from Lemma 3 (Wyner's lemma) applied to the random variable $(D, \tilde{D})$, (d) is a also a direct instance of this lemma. The inequality (b) is the independence bound on the entropy (which holds here since the random variables D and $\tilde{D}$ are discrete). The equality (c) results from the fact that D and $\tilde{D}$ are conditionally independent given X_0^T, indeed $D \in \mathcal{F}_Z^T$ whereas $\tilde{D} \in \mathcal{F}_{\tilde{Z}}^T$ and $\mathcal{F}_Z^T$ and $\mathcal{F}_{\tilde{Z}}^T$ are conditionally independent given $\mathcal{F}_X^T$. Consequently we have

$$\begin{aligned} I(X_0^T; Z_0^T, \tilde{Z}_0^T) &= \sup_{\mathcal{Q}_z, \mathcal{Q}_{\tilde{z}}} I(X_0^T; D, \tilde{D}) \\ &\leq \sup_{\mathcal{Q}_z, \mathcal{Q}_{\tilde{z}}} (I(X_0^T; D) + I(X_0^T; \tilde{D})) \\ &= I(X_0^T; Z_0^T) + I(X_0^T; \tilde{Z}_0^T). \end{aligned} \tag{61}$$

Combining the last inequality with (58) we deduce that[9]

$$I(X_0^T; \tilde{Y}_0^T) - I(X_0^T; Z_0^T) \leq I(X_0^T; \tilde{Z}_0^T). \tag{62}$$

Now using Lemma 2 we have

$$I(X_0^T; \tilde{Z}_0^T) \leq \int_0^T (\mathbb{E}[\varphi(X_t)] - \varphi(\mathbb{E}[X_t]))dt, \tag{63}$$

where

$$\varphi(x) = ((A_y - A_z)x + (\frac{A_y}{A_z} - 1)\lambda_z) \ln((A_y - A_z)x + (\frac{A_y}{A_z} - 1)\lambda_z). \tag{64}$$

[9]Note that since $\mathbb{E}\int_0^T |X_t \ln X_t| < \infty$, then $I(X_0^T; Z_0^T) < \infty$ and hence the inequality is well defined.

Notice now that

$$\varphi(x) = (\frac{A_y}{A_z} - 1)\phi_z(x) + (\frac{A_y}{A_z} - 1)\ln(\frac{A_y}{A_z} - 1)(A_z x + \lambda_z). \tag{65}$$

Plugging this identity in the inequality above, the linear term in x disappears and we are left with the inequality presented in the lemma. ■

An alternative proof of this lemma using the link provided in [24] between the MMSE and the mutual information in Poisson channels is given in Appendix B.

Theorem 3. *If* $\mathbb{E}\int_0^T |X_t \ln X_t| < \infty$*, then*

$$\frac{1}{T} I(X_0^T; Y_0^T | Z_0^T) \leq C_s \tag{66}$$

Proof: Combining (42) and the result of the two previous lemmas yields

$$I(X_0^T; Y_0^T) - I(X_0^T; Z_0^T) \leq \int_0^T (\mathbb{E}[K(X_t)] - K(\mathbb{E}[X_t]))dt, \tag{67}$$

where $K(x) = \phi_y(x) - \phi_z(x)$. A straightforward computation shows that

$$K''(x) = \frac{A_z A_y (A_y - A_z)x + \lambda_z A_y^2 - \lambda_y A_z^2}{(A_y x + \lambda_y)(A_z x + \lambda_z)}, \tag{68}$$

since $A_y \geq A_z$ and $\lambda_z A_y^2 \geq \lambda_y A_y A_z \geq \lambda_y A_z^2$ we deduce that $K''(x) \geq 0$. Moreover due to the assumption that at least one of the inequalities (5) or (6) is strict, we conclude that $K''(x) > 0$ (for $x > 0$) and $K(\cdot)$ is strictly convex.

Notice now that we have

$$\begin{aligned}
&\frac{1}{T}\int_0^T (\mathbb{E}[K(X_t)] - K(\mathbb{E}[X_t]))dt \\
&\overset{(a)}{\leq} \max_{0\leq\alpha\leq 1} \left(\max_{\rho:\int_0^1 x\rho(dx)=\alpha} \int_0^1 K(x)\rho(dx) - K(\alpha) \right) \\
&\overset{(b)}{=} \max_{0\leq\alpha\leq 1} \left(\alpha K(1) + (1-\alpha)K(0) - K(\alpha)\right),
\end{aligned} \tag{69}$$

where (a) follows from fixing $\mathbb{E}[X_t] = \alpha$ and maximizing over all distributions $\rho(x)$ on $[0,1]$ with mean α. Equality (b) follows from the convexity of $K(\cdot)$ (refer to [16] and [31]), i.e., the maximizing distribution ρ puts all the mass on the extremes $\{0,1\}$ and since the mean is α, the maximizing ρ assigns the mass α to 1 and $1-\alpha$ to 0.

The maximization of the last term shows that the optimal α^* is the unique solution to the equation

$$K'(\alpha^*) = K(1) - K(0),$$

which, after some algebraic manipulations, gives that α^* is the solution to (11). The existence of α^* follows simply from the mean value theorem, whereas the uniqueness is a consequence of the strict monotonicity of $K'(x)$.

Consequently, the following is true

$$\begin{aligned}
&\frac{1}{T}\int_0^T (\mathbb{E}[K(X_t)] - K(\mathbb{E}[X_t]))dt \\
&\leq \alpha^* K(1) + (1-\alpha^*)K(0) - K(\alpha^*) \\
&= \alpha^*(A_y - A_z) + \ln\left(\frac{\lambda_y^{\lambda_y}}{\lambda_z^{\lambda_z}}\right) + \ln\left(\frac{(A_z\alpha^* + \lambda_z)^{\lambda_z}}{(A_y\alpha^* + \lambda_y)^{\lambda_y}}\right). \qquad (70)
\end{aligned}$$

This fact when combined with (67) gives the result announced in the theorem. ■

We are now in a position to prove the converse theorem.

Theorem 4 (Converse). *If R_s is an achievable secrecy rate then $R_s \leq C_s$.*

Proof: Since the secrecy rate R_s is achievable then for all $0 < \epsilon < \frac{1}{2}$ and sufficiently large T, there exists an (M, T) code such that $\frac{\ln M}{T} \geq R_s - \epsilon$, $P_e \leq \epsilon$ and $\frac{1}{T} I(U; Z_0^T) \leq \epsilon$. Hence we have

$$\begin{aligned}
R_s &\leq \frac{\ln M}{T} + \epsilon \\
&\overset{(a)}{\leq} \frac{1}{T(1-P_e)}\left(I(X_0^T; Y_0^T | Z_0^T) + I(U; Z_0^T) + H(P_e)\right) + \epsilon \\
&\overset{(b)}{\leq} \frac{1}{1-P_e}\left(C_s + \frac{I(U; Z_0^T)}{T} + \frac{H(P_e)}{T}\right) + \epsilon \\
&\overset{(c)}{\leq} \frac{1}{1-\epsilon}\left(C_s + \epsilon + \frac{H(\epsilon)}{T}\right) + \epsilon, \qquad (71)
\end{aligned}$$

where inequality (a) follows from Lemma 4, inequality (b) from Theorem 3 and inequality (c) from the properties of the code. Now since ϵ is arbitrary, letting $\epsilon \to 0$ yields $R_s \leq C_s$. ■

V. Rate-Equivocation region

In this section we turn our attention to the rate equivocation region of the degraded Poisson wiretap channel. The level of ignorance of the eavesdropper about the transmitted message U will be measured here by the normalized

equivocation given by

$$\Delta_T = \frac{H(U|Z_0^T)}{H(U)}. \tag{72}$$

Definition A rate-equivocation pair (R, d) is said to be *achievable* for the Poisson wiretap channel if for all $\epsilon > 0$ and all sufficiently large T, there exists an (M, T) code such that

$$\begin{aligned} \frac{\ln M}{T} &\geq R - \epsilon \\ P_e &\leq \epsilon \\ \frac{H(U|Z_0^T)}{H(U)} &\geq d - \epsilon \end{aligned} \tag{73}$$

The following theorem gives the rate equivocation region for the degraded Poisson Wiretap channel.

Theorem 5. *The rate-equivocation region is the set of all rate-equivocation pairs* (R, d) *for which there exits some* $\alpha \in [0, 1]$ *such that*

$$Rd \leq \alpha \ln\left(\frac{(A_y + \lambda_y)^{A_y+\lambda_y}}{(A_z + \lambda_z)^{A_z+\lambda_z}}\right) + (1-\alpha)\ln\left(\frac{\lambda_y^{\lambda_y}}{\lambda_z^{\lambda_z}}\right) - \ln\left(\frac{(A_y\alpha + \lambda_y)^{A_y\alpha+\lambda_y}}{(A_z\alpha + \lambda_z)^{A_z\alpha+\lambda_z}}\right) \tag{74}$$

$$R \leq \alpha \ln\left((A_y + \lambda_y)^{A_y+\lambda_y}\right) + (1-\alpha)\ln\left(\lambda_y^{\lambda_y}\right) - \ln\left((A_y\alpha + \lambda_y)^{A_y\alpha+\lambda_y}\right) \tag{75}$$

$$d \leq 1 \tag{76}$$

To ease the notations, using the functions $K(\cdot)$ and $\phi_y(\cdot)$, we can rewrite the two first inequalities as $Rd \leq \alpha K(1) + (1-\alpha)K(0) - K(\alpha)$ and $R \leq \alpha\phi_y(1) + (1-\alpha)\phi_y(0) - \phi_y(\alpha)$.

Proof: The main ingredients needed to prove this theorem has been already used to obtain the secrecy capacity. More specifically, for the achievability proof we will use stochastic encoding combined with Wyner codes for the Poisson channel, and for the converse we will use the key inequality (67) established by Lemma 6 and 7.

A. Direct result

Note first that for a fixed rate R, if the rate equivocation pair (R, d) is achievable then the pair $(R, \tilde{d})$ is achievable for all $0 \leq \tilde{d} \leq d$. Hence, in order to establish the direct result, it is enough to prove that any rate-equivocation pair (R, d) satisfying $Rd = \alpha K(1) + (1-\alpha)K(0) - K(\alpha)$, $R \leq \alpha\phi_y(1) + (1-\alpha)\phi_y(0) - \phi_y(\alpha)$ and $d \leq 1$ for some $\alpha \in [0, 1]$ is achievable.

Define

$$R_u = \alpha\phi_u(1) + (1-\alpha)\phi_u(0) - \phi_u(\alpha),\ u \in \{y, z\}. \tag{77}$$

Let $\epsilon > 0$ be arbitrary (small enough) and let $R = \frac{R_y - R_z - \epsilon R}{d}$ with $d \leq 1$ and $R \leq \alpha\phi_y(1) + (1-\alpha)\phi_y(0) - \phi_y(\alpha)$. The message U to be transmitted is selected uniformly randomly from $\mathcal{U} = \{1, ..., M\}$ with $M = e^{RT}$. Define $M_y = e^{(R_y - 3\epsilon\frac{R}{2})T}$ and, following the steps described for the achievability of the secrecy capacity, construct the Wyner code $\mathcal{C} = \mathcal{W}(T, M_y, \alpha M_y)$. Partition this code arbitrarily into M smaller subcodes, i.e., $\mathcal{C} = \cup_{i=1}^{M} \mathcal{C}_i$. The cardinality of each subcode $\mathcal{C}_i$ will be equal to $M_z = \frac{M_y}{M} = e^{(R_y - R - 3\epsilon\frac{R}{2})T}$. Notice that with this choice of parameters we have

$$\frac{1}{T}\ln M_z = R_y - R - 3\epsilon\frac{R}{2} \leq R_y - Rd - 3\epsilon\frac{R}{2} = R_z - \epsilon\frac{R}{2}. \tag{78}$$

The probability of error P_e of the legitimate receiver can be made less than ϵ because the Wyner code $\mathcal{C}$ can achieve the rate R_y.

The equivocation of the code $\mathcal{C}$ can be lower bounded using the same steps used to established the upper bound on $I(U; Z_0^T)$ for the secrecy capacity, as follows

$$\Delta_T = \frac{H(U|Z_0^T)}{H(U)} = 1 - \frac{I(U; Z_0^T)}{RT} \tag{79}$$

$$\overset{(a)}{\geq} 1 - \frac{R_z}{R} + \frac{1}{RT}\ln M_z - \frac{1}{RT}\left(H(\delta) + \delta \ln M_z\right) \tag{80}$$

$$= 1 - \frac{R_z}{R} + \frac{R_y - R - 3\epsilon\frac{R}{2}}{R} - \frac{1}{RT}\left(H(\delta) + \delta \ln M_z\right) \tag{81}$$

$$\geq d - \frac{\epsilon}{2} - \frac{1}{RT}H(\delta) - \delta(\frac{R_z}{R} - \frac{\epsilon}{2}). \tag{82}$$

In the above, inequality (a) follows from (31) and $\delta = \frac{1}{M}\sum_{m=1}^{M}\delta_m$ where δ_m is the probability of error for the code $\mathcal{C}_m$ ($1 \leq m \leq M$) with the (optimal) decoder described previously.

As was discussed before, the term $\frac{1}{RT}H(\delta) + \delta(\frac{R_z}{R} - \frac{\epsilon}{2})$ can be made less than $\frac{\epsilon}{2}$ for T large enough, which means that

$$\Delta_T = \frac{H(U|Z_0^T)}{H(U)} \geq d - \epsilon. \tag{83}$$

This establishes that the rate-equivocation pair (R, d) is achievable.

B. Converse

For every (M,T) code with rate $R_T = \frac{\ln M}{T}$ and equivocation $\Delta_T = \frac{H(U|Z_0^T)}{H(U)}$ we have

$$\begin{aligned} TR_T\Delta_T = H(U|Z_0^T) &= H(U) - I(U;Z_0^T) \\ &= H(U|Y_0^T) + I(U;Y_0^T) - I(U;Z_0^T) \\ &\leq H(U|\hat{U}) + I(U;Y_0^T|Z_0^T) \\ &\leq H(P_e) + P_e \ln M + I(X_0^T;Y_0^T|Z_0^T). \end{aligned} \tag{84}$$

From Lemma 6 and 7 (cf. (67)) we have that

$$I(X_0^T;Y_0^T|Z_0^T) \leq \int_0^T (\mathbb{E}[K(X_t)] - K(\mathbb{E}[X_t]))dt. \tag{85}$$

Consequently, we deduce that

$$R_T\Delta_T \leq \frac{H(P_e) + P_e \ln M}{T} + \frac{1}{T}\int_0^T (\mathbb{E}[K(X_t)] - K(\mathbb{E}[X_t]))dt \tag{86}$$

$$\leq \frac{H(P_e) + P_e \ln M}{T} + \alpha K(1) + (1-\alpha)K(0) - K(\alpha), \tag{87}$$

with $\alpha = \frac{1}{T}\int_0^T \mathbb{E}[X_t]dt$ and the last inequality follows from the convexity of the function $K(\cdot)$. Note that since $0 \leq X_t \leq 1$ it follows that $0 \leq \alpha \leq 1$.

Similarly, we have that

$$\begin{aligned} R_T = \frac{H(U)}{T} &= \frac{1}{T}H(U|Y_0^T) + \frac{1}{T}I(U;Y_0^T) \\ &\leq \frac{1}{T}H(U|\hat{U}) + \frac{1}{T}I(X_0^T;Y_0^T) \\ &\overset{(a)}{\leq} \frac{1}{T}(H(P_e) + P_e \ln M) + \frac{1}{T}\int_0^T (\mathbb{E}[\phi_y(X_t)] - \phi_y(\mathbb{E}[X_t]))dt \\ &\overset{(b)}{\leq} \frac{H(P_e) + P_e \ln M}{T} + \alpha\phi_y(1) + (1-\alpha)\phi_y(0) - \phi_y(\alpha), \end{aligned} \tag{88}$$

where (a) follows from Fano's inequality and Lemma 2 and (b) follows from the convexity of the function $\phi_y(\cdot)$.

Assume now that (R,d) is achievable, then for all $0 < \epsilon < \frac{1}{2}$ and all sufficiently large T, there exists an (M,T) code such that $R_T \geq R - \epsilon$, $P_e \leq \epsilon$ and $\Delta_T \geq d - \epsilon$. By definition $\Delta_T \leq 1$, and hence $d \leq 1 + \epsilon$ and in light of

the previous inequalities we have

$$(R-\epsilon)(d-\epsilon) \leq \frac{H(\epsilon)+\epsilon \ln M}{T} + \alpha K(1) + (1-\alpha)K(0) - K(\alpha) \tag{89}$$

$$(R-\epsilon) \leq \frac{H(\epsilon)+\epsilon \ln M}{T} + \alpha \phi_y(1) + (1-\alpha)\phi_y(0) - \phi_y(\alpha). \tag{90}$$

Now since ϵ is arbitrary, letting $\epsilon \to 0$ yields the desired result. ■

VI. Conclusion and discussion

Motivated by the practical advantages of optical communication over RF for secure communication, we have derived the secrecy capacity and characterized the rate-equivocation region of the degraded Poisson wiretap channel.

Several interesting problems remain open and deserve further investigation. One is the non-degraded Poisson Wiretap channel. One can imagine a situation in which the eavesdropper is equipped with a powerful detector characterized by a negligible dark current (i.e., $\lambda_z = 0$). If the detector of the legitimate receiver has a higher received power from the transmitter but is more prone to dark current, then the channel will not be degraded. This is a practically-important situation but is not covered by the results of this paper.

Another issue that we have not considered is fading. Indeed, for wireless optical communications, atmospheric turbulence can induce random fluctuations of the intensity of the transmitted light beam [23], which creates fading and complicates secure communication. Note, however, that this fading is fundamentally different from multipath fading and is more manageable from the standpoint of achieving secure communication.

MIMO Poisson channels have received some interest lately (see [32] and the references therein), and as has been done in the Gaussian setting, it would be interesting to see the impact of having multiple antennas on the secrecy capacity in the Poisson regime.

We believe that the results derived in this paper and the tools used to derive them could be used to address these problems.

APPENDIX A

AN MMSE PROOF FOR LEMMA 6

In this first appendix, we provide an alternative proof for Lemma 6. This proof uses the link established in [24] between the MMSE and the mutual information in Poisson channels. Note first that since $\mathbb{E}\int_0^T |X_t \ln X_t| < \infty$, we have that $I(X_0^T; \mathcal{P}_0^T(A_y X_0^T + \lambda))$ is differentiable and Theorem 3 in [24] states that

$$\frac{d}{d\lambda} I(X_0^T; \mathcal{P}_0^T(A_y X_0^T + \lambda)) = \int_0^T \mathbb{E}\{\ln(A_y X_t + \lambda) - \ln\langle A_y X_t + \lambda\rangle_T\} dt. \tag{91}$$

Notice now that

$$I(X_0^T; \tilde{Y}_0^T) - I(X_0^T; Y_0^T) = \int_{\lambda_y}^{\lambda_y + \tilde{\lambda}} \frac{d}{d\lambda} I(X_0^T; \mathcal{P}_0^T(A_y X_0^T + \lambda)) d\lambda. \tag{92}$$

Therefore

$$I(X_0^T; Y_0^T) - I(X_0^T; \tilde{Y}_0^T) = \int_{\lambda_y}^{\lambda_y + \tilde{\lambda}} \left(\int_0^T (\mathbb{E}\{\ln\langle A_y X_t + \lambda\rangle_T\} - \mathbb{E}\{\ln(A_y X_t + \lambda)\}) dt \right) d\lambda. \tag{93}$$

Since the function $\ln(\cdot)$ is concave, using Jensen's inequality and the iterative conditioning property we have

$$\mathbb{E}\{\ln\langle A_y X_t + \lambda\rangle_T\} \leq \ln \mathbb{E}[\langle A_y X_t + \lambda\rangle_T] = \ln(A_y \mathbb{E}[X_t] + \lambda).$$

Making use of this inequality and the fact that $\lambda_y + \tilde{\lambda} = \frac{A_y}{A_z}\lambda_z$ we deduce that

$$I(X_0^T; Y_0^T) - I(X_0^T; \tilde{Y}_0^T) \leq \int_0^T \left(\int_{\lambda_y}^{\frac{A_y}{A_z}\lambda_z} \ln(A_y \mathbb{E}[X_t] + \lambda) d\lambda - \mathbb{E}\{ \int_{\lambda_y}^{\frac{A_y}{A_z}\lambda_z} \ln(A_y X_t + \lambda) d\lambda \} \right) dt, \tag{94}$$

where we have also invoked Fubini's theorem to make the necessary exchanges between the integrals and the expectation operator. The desired inequality is then obtained after some algebraic manipulations using the elementary identity

$$\int \ln(A_y x + \lambda) d\lambda = (A_y x + \lambda) \ln(A_y x + \lambda) - \lambda. \tag{95}$$

APPENDIX B

AN MMSE PROOF FOR LEMMA 7

Here we provide an alternative proof for Lemma 7. For ease of notations define $W_t = A_y X_t + \frac{A_y}{A_z}\lambda_z$. Using Theorem 4 in [24] we obtain that

$$\begin{aligned}\frac{d}{d\alpha} I(W_0^T; \mathcal{P}_0^T(\alpha W_0^T)) &= \int_0^T \mathbb{E}[W_t \ln(\alpha W_t)]dt - \int_0^T \mathbb{E}[\mathbb{E}[W_t|\mathcal{P}_0^T(\alpha W_0^T)] \ln(\mathbb{E}[\alpha W_t|\mathcal{P}_0^T(\alpha W_0^T)])]dt \\ &= \int_0^T \mathbb{E}[W_t \ln W_t]dt - \int_0^T \mathbb{E}[\mathbb{E}[W_t|\mathcal{P}_0^T(\alpha W_0^T)] \ln(\mathbb{E}[W_t|\mathcal{P}_0^T(\alpha W_0^T)])]dt, \end{aligned} \tag{96}$$

where the second equality is obtained after some simplifications using the identity $\mathbb{E}[\mathbb{E}[W_t|\mathcal{P}_0^T(\alpha W_0^T)]] = \mathbb{E}[W_t]$. Now by the convexity of the function $C(x) = x\ln(x)$, Jensen's inequality gives

$$\begin{aligned}\mathbb{E}[C(\mathbb{E}[W_t|\mathcal{P}_0^T(\alpha W_0^T)])] &\geq C(\mathbb{E}[\mathbb{E}[W_t|\mathcal{P}_0^T(\alpha W_0^T)]]) \\ &= C(\mathbb{E}[W_t]). \end{aligned} \tag{97}$$

It follows therefore that

$$\frac{d}{d\alpha} I(W_0^T; \mathcal{P}_0^T(\alpha W_0^T)) \leq \int_0^T \mathbb{E}[W_t \ln W_t]dt - \int_0^T \mathbb{E}[W_t]\ln \mathbb{E}[W_t]dt. \tag{98}$$

Clearly we have that $I(W_0^T; \mathcal{P}_0^T(\alpha W_0^T)) = I(X_0^T; \mathcal{P}_0^T(\alpha W_0^T))$. Also we have that $\tilde{Y}_0^T = \mathcal{P}_0^T(W_0^T)$ and $Z_0^T = \mathcal{P}_0^T(\frac{A_z}{A_y} W_0^T)$. Consequently

$$I(X_0^T; \tilde{Y}_0^T) - I(X_0^T; Z_0^T) = \int_{\frac{A_z}{A_y}}^1 \frac{d}{d\alpha} I(W_0^T; \mathcal{P}_0^T(\alpha W_0^T))d\alpha. \tag{99}$$

Using the previous inequality, we conclude that

$$\begin{aligned} I(X_0^T; \tilde{Y}_0^T) - I(X_0^T; Z_0^T) &\leq \int_{\frac{A_z}{A_y}}^1 \left(\int_0^T \mathbb{E}[W_t \ln W_t]dt - \int_0^T \mathbb{E}[W_t]\ln \mathbb{E}[W_t]dt \right) d\alpha \\ &= (1 - \frac{A_z}{A_y}) \left(\int_0^T \mathbb{E}[(A_y X_t + \frac{A_y}{A_z}\lambda_z)\ln(A_y X_t + \frac{A_y}{A_z}\lambda_z)]dt \right. \\ &\quad \left. - \int_0^T (A_y \mathbb{E}[X_t] + \frac{A_y}{A_z}\lambda_z)\ln(A_y \mathbb{E}[X_t] + \frac{A_y}{A_z}\lambda_z)dt \right). \end{aligned} \tag{100}$$

After some simplifications, the last inequality gives the desired result, i.e.,

$$I(X_0^T; \tilde{Y}_0^T) - I(X_0^T; Z_0^T) \leq (\frac{A_y}{A_z} - 1) \int_0^T (\mathbb{E}[\phi_z(X_t)] - \phi_z(E[X_t]))\, dt. \tag{101}$$

REFERENCES

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, pp. 656-715, October 1949.

[2] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, pp. 1355-1387, October 1975.

[3] I. Csisár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339-348, May 1978.

[4] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 24, no. 4, pp. 451-456, July 1978.

[5] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas: The MIMOME channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547-2553, June 2009.

[6] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," in *Proc. IEEE Int. Symp. Information Theory*, Toronto, ON, Canada, pp. 524-528., July 2008.

[7] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multi-antenna wiretap channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547-2553, June 2009.

[8] P. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4687-4698, October 2008.

[9] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2453-2469, June 2008.

[10] J. R. Barry, "Wireless Infrared Communications", *Kluwer Academic Publishers*, Boston, MA, 1994.

[11] J. M. Kahn and J. R. Barry, "Wireless infrared communications," *Proceedings of the IEEE*, vol. 85, pp. 265-298, Feburary 1997.

[12] R. J. Green, "Secure communications: the Infrared alternative," in *ICTON Mediterranean Winter Conference*, 2007.

[13] Z. Xu and B. M. Sadler, "Ultraviolet communications: Potential and state-of-the-art," *IEEE Communications Magazine*, pp. 67-73, May 2008.

[14] S. Adee, "Ultraviolet radios beam to life," *IEEE Spectrum*, May 2009.

[15] V. G. Sidorovich, "Optical countermeasures and security of free-space optical communication links," in *Proc. of the SPIE: Advanced Free-Space Optical Communications Techniques and Technologies*, vol. 5614, pp. 97-108, October 2004.

[16] Y. Kabanov, "The capacity of a channel of the Poisson type," *Theory of Probability and its Appl.*, vol. 23, pp. 143-147, 1978.

[17] M. H. A. Davis, "Capacity and cutoff rate for Poisson-type channels," *IEEE Trans. Inf. Theory*, vol. 26, pp. 710-715, November 1980.

[18] A. D. Wyner, "Capacity and error exponent for the direct detection photon channel: Part I," *IEEE Trans. Inf. Theory*, vol. 34, no. 6, pp. 1449-1461, November 1988.

[19] A. Lapidoth and S. Shamai (Shitz), "The Poisson multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 44, pp. 488-502, March 1998.

[20] S. I. Bross, M. V. Burnashev and S. Shamai (Shitz), "Error exponents for the two-user Poisson multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 47, no. 5, pp. 1999-2016, July 2001.

[21] A. Lapidoth, E. Telatar and R. Urbanke, “On wide-band broadcast channels”, *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3250-3258, December 2003.

[22] A. Sokolovsky and S.I. Bross “Attainable error exponents for the Poisson broadcast channel with degraded message sets,” *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 364-374, January 2005.

[23] K. Chakraborty and P. Narayan, “The Poisson fading channel,” *IEEE Trans. Inf. Theory*, vol. 53, no. 7, pp. 2349-2364, July 2007.

[24] D. Guo, S. Shamai (Shitz) and S. Verdú, “Mutual information and conditional mean estimation in Poisson channels”, *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1837-1849, May 2008.

[25] D. Guo, S. Shamai (Shitz), and S. Verdú, “Mutual information and minimum mean-square error in Gaussian channels,” *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1261-1283, April 2005.

[26] E. Ekrem, and S. Ulukus “Secrecy Capacity Region of the Gaussian Multi-Receiver Wiretap Channel,” in *Proc. of IEEE International Symposium on Information Theory*, Seoul, Korea, June 2009.

[27] R. Bustin, R. Liu, H. V. Poor and S. Shamai (Shitz), “An MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel,” *EURASIP Journal on Wireless Communications and Networking*, 2009.

[28] M. Van Dijk, “On a special class of broadcast channels with confidential messages,” *IEEE Trans. Inf. Theory*, vol. 43, no.2, pp. 712-714, Mars 1997.

[29] A.D. Wyner, “A definition of conditional mutual information for arbitrary ensembles”, *Inform. Contr.*, vol. 38, pp. 51-59, 1978.

[30] P. Brémaud, Point Processes and queues: Martingale Dynamics, Springer-Verlag, New York, 1981.

[31] M. Krein and A. Nudelman, “The Markov moment problem and extremal problems,” *Translations of Mathematical Monographs Providence*, RI: Amer. Math. Soc., vol. 5, 1977.

[32] K. Chakraborty, S. Dey, M. Franceschetti, “Outage capacity of MIMO Poisson fading channels,” *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 4887-4907, November 2008.